

# Shearlet-based vs. Photometric-based Visual Servoing for Robot-assisted Medical Applications

Lesley-Ann Duflot, Alexandre Krupa, Brahim Tamadazte, Nicolas Andreff

## ► To cite this version:

Lesley-Ann Duflot, Alexandre Krupa, Brahim Tamadazte, Nicolas Andreff. Shearlet-based vs. Photometric-based Visual Servoing for Robot-assisted Medical Applications. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'16, Oct 2016, Daejeon, South Korea. pp.4099-4104. hal-01355414

**HAL Id: hal-01355414**

**<https://hal.inria.fr/hal-01355414>**

Submitted on 23 Aug 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Shearlet-based vs. Photometric-based Visual Servoing for Robot-assisted Medical Applications

Lesley-Ann Duflot<sup>1,2</sup>, Alexandre Krupa<sup>1</sup>, Brahim Tamadazte<sup>2</sup>, and Nicolas Andreff<sup>2</sup>

**Abstract**—This paper deals with the development of a vision-based controller for robot-assisted medical applications. It concerns the use of shearlet coefficients in case of ultrasounds (US) images as visual signal inputs and the design of the associated interaction matrix. The proposed controller was validated in both simulation and on an experimental test bench which consists of a robotic arm holding an US probe in contact with a realistic abdominal phantom. Also, the proposed control scheme was compared to the photometry-based visual servoing approach in order to evaluate its efficiency in different conditions of use (nominal and unfavorable conditions).

## I. INTRODUCTION

Shearlet transform is a recent mathematical tool, introduced in 2005, for signal representation [1]. It is considered as a natural extension of the wavelet transform. The shearlets provide a better sensitivity to edge and anisotropic image features than wavelets. Throughout the last decade, several interesting applications of shearlets have emerged such as image inpainting [2], denoising [3], edge detection [4], speckle filtering on US images [5], etc. More recently, we have developed a new ultrasound-based visual servoing scheme for medical applications in which the visual signal inputs were the subsampled shearlet coefficients [6].

US images are widely used for non-invasive examinations of internal organs. Real-time US images can be also used as visual feedback for guiding mini-invasive surgical procedures. Unfortunately, medical ultrasonography implies high cognitive demands on the surgery. It is for this reason that using an US probe held by a robotic manipulator is very helpful. A robotic system during an US examination can remove surgeon tremor and gives a better precision of the surgical gesture.

Since few years, different vision-based controllers were reported in the literature with the aim to control an US probe held by a robot. Therefore, several visual servoing approaches using visual information extracted from the US images were proposed [7], [9], [10]. For instance, the objectives of these approaches are: to maintain the visibility of an organ during the echography examination [11]; to provide an automatic positioning of the probe in order to retrieve a pre-operative image [7]; or to automatically compensate the

physiologic motions (e.g., breathing) in order to stabilize US images during an examination [9].

The efficiency of a visual servoing approach depends on the choice of the appropriate visual features. It can be points [11], lines, moments [7], wavelets [8], etc. Generally, visual servoing schemes deal with *visual tracking* algorithms and their success is directly related on the ability to detect, match and track over time the visual features in the US images. To overcome these limitations, original visual servoing approaches have been introduced showing that the design of an image-based visual controller can totally remove the visual tracking process [12], [13]. The techniques are called direct visual servoing.

In [9], the authors proposed an alternative 6 DOF visual servoing scheme by considering directly as visual inputs the pixel intensity of the US images. This method has proven its reliability in terms of robustness and accuracy thanks to the visual information redundancy. Also, in [6], it was reported a preliminary study concerning the design of a new controller that uses subsampled shearlet coefficients instead of pure image signal information. Among the various advantages of this method, using shearlet coefficients in the control loop allows selecting a set of *noiseless* and redundant features more robust than pixel intensities. In this paper, a comparison study will demonstrate that the shearlet-based visual servoing can be more accurate and robust with respect to the intensity-based framework. In addition, our controller was experimentally tested in several realistic conditions of use i.e., favorable conditions and unfavorable ones with Gaussian or speckle noises. The latter are the main drawback of US images. The results demonstrate the efficiency of the proposed controller in terms of accuracy, convergence, robustness and repeatability.

In the sequel, Section II reviews the general basics of the photometric and shearlet-based visual servoing. Both approaches are compared in Section III where the cost-functions were computed and studied. Section IV and Section V show the simulation and the experimental results.

## II. US-BASED VISUAL SERVOING FRAMEWORKS

### A. Photometry-based Method

According to [14], the aim of a visual servoing scheme is to control the motion of a robot in order to allow a set of visual signal  $\mathbf{s}$  ( $\mathbf{s} \in \mathbb{R}^k$ ) defining a robot pose  $\mathbf{r}(t) \in SE(3)$  (i.e.,  $\mathbf{s} = \mathbf{s}(\mathbf{r}(t))$ ) to reach a set of desired signal  $\mathbf{s}^*$  ( $\mathbf{s}^* \in \mathbb{R}^k$ ) by minimizing a visual error given by:

$$\mathbf{e} = \mathbf{s}(\mathbf{r}(t)) - \mathbf{s}^* \quad (1)$$

<sup>1</sup>L. A. Duflot and A. Krupa are with the Inria Rennes-Bretagne Atlantique, Campus universitaire de Beaulieu, Rennes 35042, France  
lesley-ann.duflot@inria.fr - alexandre.krupa@inria.fr

<sup>2</sup>L. A. Duflot, B. Tamadazte and N. Andreff are with FEMTO-ST, AS2M, Univ. Bourgogne Franche-Comté, Univ. de Franche-Comté/CNRS/ENSMM, 25000 Besançon, France  
brahim.tamadazte@femto-st.fr - nicolas.andreff@femto-st.fr

The variation of  $\mathbf{s}$  is linked to the velocity screw vector  $\mathbf{v} = (v_x \ v_y \ v_z \ \omega_x \ \omega_y \ \omega_z)^T$  of the US probe frame by  $\dot{\mathbf{s}} = \mathbf{L}_s \mathbf{v}$  where  $\mathbf{L}_s \in \mathbb{R}^{k \times 6}$  is called *interaction matrix*.

In order to obtain an exponential decrease of the visual error, the Gauss-Newton method is usually used [14]:

$$\mathbf{v} = -\lambda \mathbf{L}_s^+ (\mathbf{s}(t) - \mathbf{s}^*) \quad (2)$$

where  $\lambda$  is a positive gain and  $\mathbf{L}_s^+$  is the Moore-Penrose pseudo-inverse of the interaction matrix  $\mathbf{L}_s$ .

In [9], an US-based visual servoing, using pixel intensity as visual features was developed. In this case, by considering an image  $I_{i,j}$ , the visual feature vector  $\mathbf{s}$  is given by:

$$\mathbf{s} = (I_{1,1}, I_{1,2}, \dots, I_{m,n}) \quad (3)$$

where  $I_{i,j}$  is the pixel value at the coordinates  $i, j$  and  $m, n$  the image size.

For a sake of simplicity, in the following, we will consider only the control of the 3 DOF of the probe lying in the plane of the 2D US image that corresponds to the probe velocity vector  $\mathbf{v} = (v_x \ v_y \ \omega_\theta)^T$  with  $\omega_\theta = \omega_z$  being the angular velocity around the orthogonal axis of the US image plane. Therefore, each line  $\mathbf{L}_{I_{i,j}} \in \mathbb{R}^{1 \times 3}$  of the interaction matrix  $\mathbf{L}_s \in \mathbb{R}^{(m,n) \times 3}$  that links the variation of the intensity of a pixel  $(i, j)$  to this in-plane velocity vector  $\mathbf{v}$  is given by [9]:

$$\mathbf{L}_{I_{i,j}} = [\nabla I_x(i, j) \ \nabla I_y(i, j) \ x \nabla I_y(i, j) - y \nabla I_x(i, j)] \quad (4)$$

where  $x$  and  $y$  are the meter 2D coordinates of the pixel  $(i, j)$  expressed in a 2D frame centered in the image such that:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} s_x(j - j_0) \\ s_y(i - i_0) \end{pmatrix} \quad (5)$$

with  $(s_x, s_y)$  the size in meter of one pixel and  $(i_0, j_0)$  the pixel coordinates of the center of the 2D US image.

Here,  $\nabla I_{i,j} = [\nabla I_x(i, j) \ \nabla I_y(i, j)]$  corresponds to the 2D image gradient associated to the pixel  $(i, j)$  that can be computed with a simple Sobel filter. The interaction matrix related to the visual features  $\mathbf{s}$  is then obtained by stacking the matrices  $\mathbf{L}_{I_{i,j}}$  related to each pixel intensity:

$$\mathbf{L}_s = \begin{bmatrix} \mathbf{L}_{I_{1,1}} \\ \vdots \\ \mathbf{L}_{I_{m,n}} \end{bmatrix} \quad (6)$$

### B. Shearlet-based Method

Very similar to a wavelet representation, the shearlet transform of an image allows a multi-scale analysis by shearing the frequency domain of this image with different scales  $\alpha$  and orientations  $\beta$ . This shearing of the frequency domain with different orientations is the major difference between wavelets and shearlets and explains the better sensibility of the latter, for example, to edge detection [4]. Fig. 1 illustrates the tilting of the shearlet transform of an image  $I$  in the frequency domain with 4 tilting directions at a scale  $\alpha = 1$  and 8 tilting directions at a scale  $\alpha = 2$ .

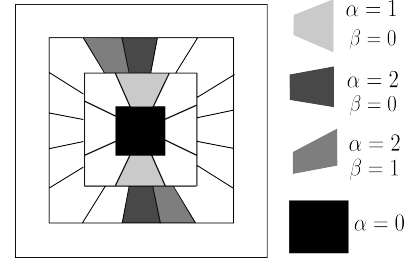


Fig. 1. Representation of a frequency plane divided by the shearlet transform.

Thereby, the shearlet transform of an image  $I \in L^2(\mathbb{R}^2)$  is given by  $SH_\phi I(\alpha, \beta)$  as:

$$I \rightarrow SH_\phi I(\alpha, \beta) = \langle I, \phi_{\alpha, \beta, t} \rangle \quad (7)$$

where the convolution function set is defined as:

$$\left\{ \phi_{(\alpha, \beta, t)} = T_t D_{A_\alpha} D_{S_\beta} \phi, \ \alpha \in \mathbb{R}, \ \beta \in \mathbb{R}, \ t \in \mathbb{R}^2 \right\} \quad (8)$$

where  $T_t$  is a translation operator,  $D_{A_\alpha}$  is a dilatation operator,  $D_{S_\beta}$  is a shearing operator and  $\phi$  is a generating function. For more details about shearlet computation, the reader is invited to refer to [15].

In our previous work [6], we developed a preliminary version of the shearlet-based visual servoing using subsampled shearlet coefficients [3], where an image of  $n \times m$  pixels is represented by  $n \times m$  coefficients. The latter were ranged in function of scale  $\alpha$  and orientation  $\beta$  almost similar to the wavelet representation. The visual servoing framework considered as visual features  $\mathbf{s}$ , the  $k$  coefficients of the level  $\alpha = 1$ , with  $k \ll n \times m$  due to the subsampling of the used shearlet transform version.



Fig. 2. Computed shearlet coefficients of a head phantom (first image) in case of  $\alpha = 0$  (second image) and  $\alpha = 1$  for the four other ones.

However, in this paper, we choose a more efficient and adequate shearlet implementation. It consists of a *non-subsampled* representation [16]. Thus, for each couple  $(\alpha, \beta)$ , we have now  $n \times m$  coefficients in shearlet domain that represent an image as shown in Fig. 2. Then, in the current work, we propose to consider the pixel values of this new image modality in the same way that the photometric approach presented in Section II-A but by replacing the image photometric intensity with the shearlet coefficient values. Moreover, we consider not only one image as visual information but 5 images (Fig. 2) that represent the  $5 \times m \times n$  coarsest shearlet coefficients corresponding to the scales  $\alpha = 0$  and  $\alpha = 1$  and we define therefore the set of visual features as follows:

$$\mathbf{s} = (SH_\phi I(0, 0), SH_\phi I(1, 0), SH_\phi I(1, 1), SH_\phi I(1, 2), SH_\phi I(1, 3)) \quad (9)$$

with  $\mathbf{s}$  being now a visual feature column vector of length  $5 \times m \times n$  that stacks the values of the pixels of the 5 shearlet coefficient images.

The interaction matrix  $\mathbf{L}_s \in \mathbb{R}^{(5.m.n) \times 3}$  related to this new shearlet-based visual feature vector  $\mathbf{s}$  is obtained by stacking the interaction matrices given by (6) computed for each image  $SH_\phi \mathbf{I}(\alpha, \beta)$ :

$$\mathbf{L}_s = \begin{bmatrix} \mathbf{L}_{SH_\phi \mathbf{I}(0,0)}, \mathbf{L}_{SH_\phi \mathbf{I}(1,0)}, \mathbf{L}_{SH_\phi \mathbf{I}(1,1)}, \mathbf{L}_{SH_\phi \mathbf{I}(1,2)}, \mathbf{L}_{SH_\phi \mathbf{I}(1,3)} \end{bmatrix} \quad (10)$$

Then the classical control law (2) is applied to minimize the visual error  $\mathbf{s} - \mathbf{s}^*$ . In this paper, we are able to compute the analytic interaction matrix (10) in contrary to [6] in which only a numerical estimated interaction matrix was developed.

### III. COST-FUNCTIONS STUDY

In order to evaluate the efficiency of the proposed controller in term of convergence domain, we have computed the cost-function shape when simulating displacements on the probe (translations along  $x$  and  $y$  axes and rotation  $\theta$  around the orthogonal axis of the image plane) with respect to the initial position. Our cost-function is given by:

$$C(\mathbf{s}) = \left( ((\mathbf{s} - \mathbf{s}^*)^\top (\mathbf{s} - \mathbf{s}^*)) / N_{pix} \right)^{\frac{1}{2}} \quad (11)$$

where,  $N_{pix}$  is the number of pixel in the initial image.

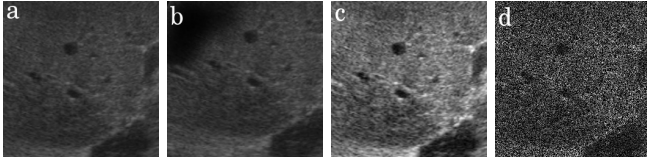


Fig. 3. (a) nominal case, (b) partial occlusion, (c) gaussian noise and (d) speckle noise.

In this section, is provided a comparison between the reliability of the photometric-based visual servoing (3) *versus* the shearlet-based one (9), this in both nominal and unfavorable conditions of use. The validation task consists using of: 1) a desired image grabbed in nominal conditions (Fig.3(a)), 2) an image with partial occlusions representing nearly one quarter of the region of interest (ROI) (Fig.3(b)), 3) an image where we applied a Gaussian noise of mean 0 and variance of 0.05 (Fig.3(c)) and 4) an image with a speckle noise i.e., the pixel intensities were multiplied with a random value of an amplitude of 1 and a variance of 0.8 (Fig.3(d)). Therefore, Fig. 4(a), (b),

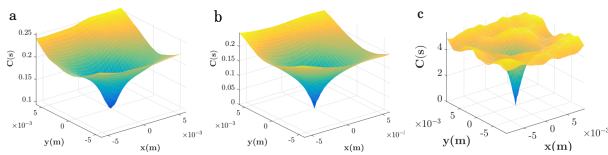


Fig. 4.  $x-y$  translations cost-function in nominal case: (a) photometric-based, (b) non-subsampled shearlet-based and (c) subsampled shearlet-based.

and (c) depict the computed, in nominal conditions, cost-function shapes for photometric-based, non-subsampled and subsampled shearlet-based methods, respectively. As can be highlighted, each cost-function has a well-defined global minimum as well as a wide top and a tight bottom. It can be considered that the three methods are almost similar in nominal conditions except the subsampled shearlet method, using a numerical estimated interaction matrix, which has lower convergence domain.

In addition, these cost-functions were computed again for each approach in unfavorable conditions of use i.e., partial occlusions and under Gaussian noise. Again, in these cases, the methods present almost similar curves (i.e., a well-defined global minimum). However, when speckle noise is applied to the images (US images are usually affected by this kind of noise), we obtain respectively the cost-function shape depicted on Fig. 5(a) for the intensity-based approach and respectively the ones presented on Fig. 5(b) and Fig. 5(c) for the non-subsampled shearlet-based and subsampled shearlet-based frameworks. The shearlet method gives better conditions of working with respect to the intensity one since their cost-function shapes are smoother and not affected by noise.

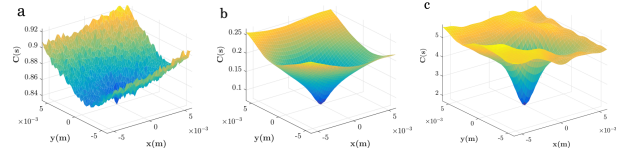


Fig. 5.  $x-y$  translations cost-function with speckle: (a) photometric-based, (b) non-subsampled shearlet-based and (c) subsampled shearlet-based.

### IV. SIMULATION VALIDATION

The approach was tested using a developed C++ simulator which is very appropriate to validate the theoretical developments in different conditions avoiding the risk of damage to the experimental setup. The simulator allows the motion control of a virtual 2D US probe that interacts with a 3D ultrasound volume of a human liver that was previously acquired with a real probe. It was implemented using the: open-source Visual Servoing Platform (ViSP) ([www.irisa.fr/lagadic/visp](http://www.irisa.fr/lagadic/visp)), Visualization ToolKit library (VTK) ([www.vtk.org](http://www.vtk.org)) and Shearlab3D library ([www.shearlab.org](http://www.shearlab.org)) in MATLAB. In order to compare carefully the photometric and shearlet methods, all the computations were performed using the same conditions and the same material.

Otherwise, the test consists of performing a 3 DOF automatic positioning task using the proposed visual controller and comparing it with the same positioning task using photometric-based approach. To do this, the task consists of reaching a desired position (i.e., desired image) from an arbitrary initial one. The control velocities  $\mathbf{v} = (v_x \ v_y \ \omega_\theta)^T$  computed by the controller (2) with  $\mathbf{s}$  and  $\mathbf{L}_s$  given by (9) and (10) respectively are applied to the virtual probe in order to reach automatically the desired position.

We proposed to perform 6 different tests to compare the non-subsampled shearlet and the photometric approaches

where 3 cases were considered: nominal conditions, addition of Gaussian noise and application of speckle noise. Fig. 6 shows the initial image (a) and the image difference  $I_{diff}$  (b) that are observed at the beginning of each of the 6 tests. This image difference between the current  $I$  and desired  $I^*$  images is computed from:

$$I_{diff} = \frac{(I - I^*) + 255}{2} \quad (12)$$

For each test, the initial positioning error was set to  $\Delta \mathbf{e}_{init} = (e_x = 10, e_y = 5, e_\theta = 20)$  (translations along  $x$  and  $y$  axis and rotation  $\theta$  in the image plane expressed in mm and degrees) and the control gain  $\lambda$  was tuned empirically to allow the convergence of the system as fast as possible without generating oscillations at the end of the task. The ROI of the desired image is surrounded in cyan (Fig. 7).

The period of the control loop was fixed to 40 ms in the simulation corresponding therefore to a frame rate of 25 images/second. For both approaches, the best performances were obtained with a gain  $\lambda = 0.08$  in nominal case,  $\lambda = 0.03$  for Gaussian noise, and  $\lambda = 0.05$  for speckle noise.

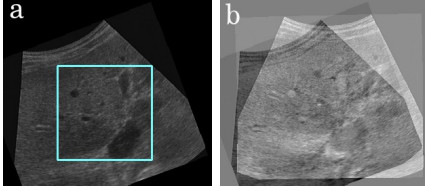


Fig. 6. (a) initial image, (b) initial error  $I_{diff}$  for all the tests.

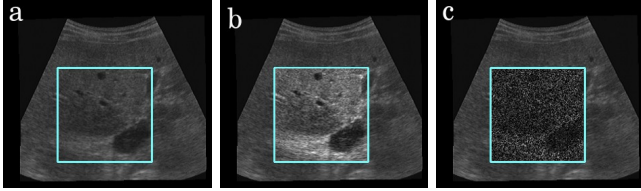


Fig. 7. Desired image in (a) nominal conditions (b) Gaussian noise (c) Speckle noise with the ROI surrounded in cyan color.

#### A. Nominal Conditions

In nominal conditions, the final error obtained with the shearlet approach and with the photometric approach converges both to zero. The positioning task succeeds in both cases.

Fig. 8 presents the positioning error and the velocity evolution of the shearlet approach in the top and the positioning error and the velocity evolution of the photometric approach in the bottom. One can observe that the velocity decreases gradually with the shearlet approach contrary to the photometric one that blows up just before the reach of the desired position. Moreover, the most striking aspect to observe is the convergence duration of the positioning task: nearly 1.2 seconds for the shearlet approach against nearly 10.8 seconds for the photometric one. It is also important to notice that increasing the gain  $\lambda$  would induce oscillations at the end of the task for both methods.

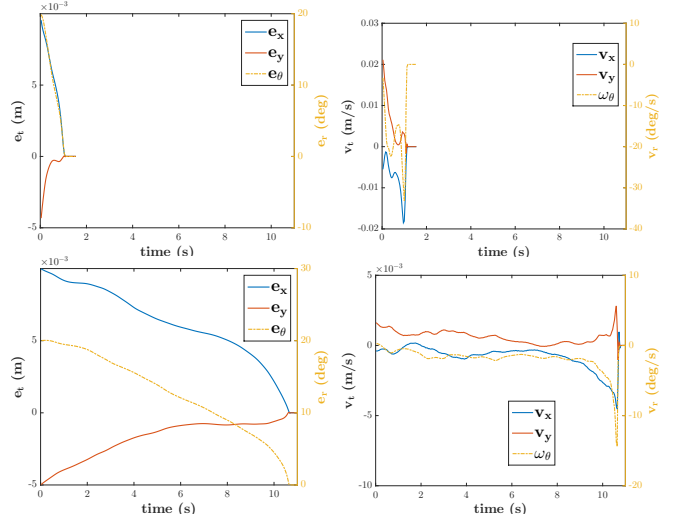


Fig. 8. Nominal conditions: (left) positioning error decay in each DOF vs. time, and (right) velocity evolution for (top) shearlet approach and (bottom) photometric approach.

#### B. Under Gaussian Noise

Under Gaussian noise, the final error with the shearlet approach and with the photometric method converge near to zero. The positioning task succeeds in both cases. The figures representing the positioning error and the velocity evolution of both methods are not represented because they are very similar in terms of shape and of convergence time to those of Fig. 8. Here, again we can conclude that the shearlet approach is definitively faster than the photometric one.

#### C. Under Speckle Noise

In the case, where we apply a speckle noise, the final error with the shearlet approach (Fig. 9 (b)) converges near zero in opposite to the photometric controller (Fig. 9 (c)). In fact, this demonstrates that photometric-based method is less robust in presence of speckle noise.

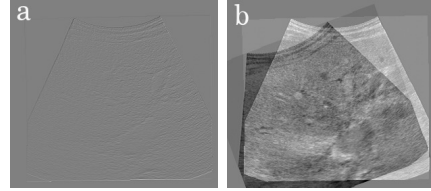


Fig. 9. Speckle noise: (a) final error for shearlet approach and (b) final error for photometric one.

The positioning error decay and the velocity evolution obtained for the shearlet approach (Fig. 10 top) are smooth and the convergence time is closed to 1.2 seconds as it was the case in nominal conditions. In opposite the photometric approach shows an unstable behaviour (Fig.10 bottom) and unfortunately converges to a local minimum. The cost-function shape study of Section III predicted the faster convergence of the shearlet approach and the simulations



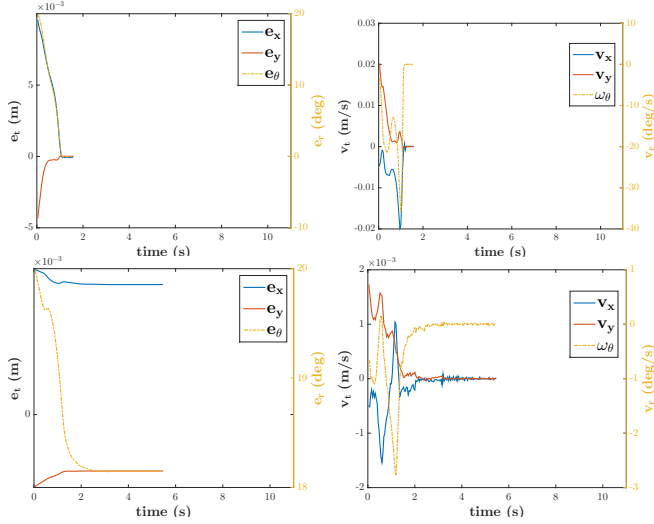


Fig. 10. Speckle noise: (left) error decay in each DOF vs. time, and (right) velocity evolution for (top) shearlet approach and (bottom) photometric approach.

validate this aspect. Moreover, from this simulation tests we can conclude that the shearlet approach converges with a decay rate 10 times faster than the photometric method. Even more, the speckle cost-function (Fig. 5) announced a potential difficulty of the photometric method to converge under speckle noise and the simulation verified this aspect. These preliminary simulation results demonstrate the robustness of the shearlet approach in three different cases. In the next section, we will also consider the presence of image occlusions during real experiments where shadows occur systematically.

## V. EXPERIMENTAL VALIDATION

The proposed controller was tested and validated in experimental conditions using a realistic test bench (Fig. 11) equipped with a 6 DOF anthropomorphic robotic arm of type of *Adept Viper s850*; a 2-5MHz 2D-US transducer (*C60, Sonosite 180+*) providing  $480 \times 640$  pixels US images (pixel size of  $0.35 \times 0.36\text{mm}^2$ ) with a depth of 12cm; a 6 axes force sensor *ATI Gamma SI-65-5*; a 3 GHz PC running under a Linux distribution in which were implemented, in C++, both developed approaches and control communication between the PC and the robot and an abdominal phantom of type of *AB-41900-030 Kyoto Kagaku-ABDFAN*.

The objective of this experimental test is to reproduce the simulation tests on a realistic setup closer to clinical US-based examination scenari. For validation purpose, the US probe attached to the 6 DOF *Viper* robot is positioned to a reference location (preoperative image) where the desired shearlet coefficients were computed, then without moving the phantom, we position randomly the probe at a new location which is considered as the initial position (intraoperative image). Thereby, the objective of controller is to retrieve automatically and accurately the preoperative image by controlling the robot displacements.

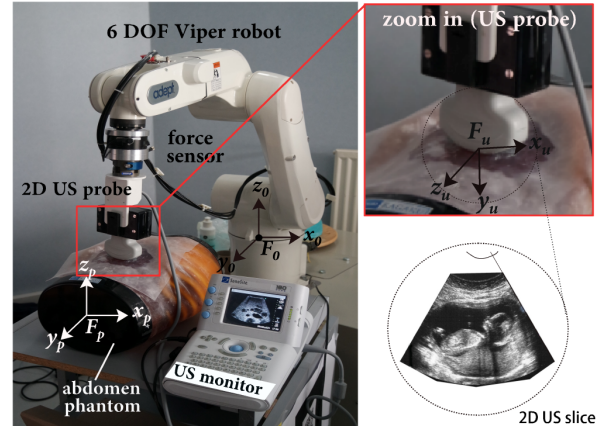


Fig. 11. Photography of the experimental set-up.

Furthermore, the y-translation corresponds to the contact direction between the US probe and the phantom. Therefore, we propose to apply a constant pressure force along this direction to maintain a good image quality. This was done by using a force-based controller on the y-translation instead of the visual controller (for more details, please refer to [9]). The aim is to maintain a desired contact force of 2N while the others 2 DOF are controlled by our shearlet-based control law.

Some images grabbed during the positioning task that exhibit a lot of shadows are presented in Fig. 12. They show that the shearlet approach can deal also with partial occlusions.

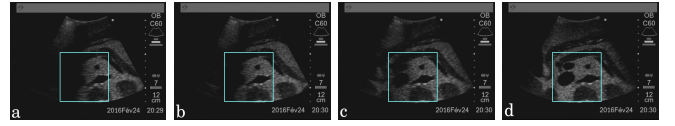


Fig. 12. Image sequence grabbed during 3 DOF positioning task using the shearlet-based control law. (a) current image at  $t = 0$ , (b) at  $t = 14\text{s}$ , (c) at  $t = 20\text{s}$  and (d) at  $t = 26\text{s}$ .

Fig. 13 (left) depicts the smooth positioning error decay during the task achievement and Fig. 13 (right) shows the robot velocities evolution. One can see that the controller behavior in experimental conditions remains very similar to the simulation ones.

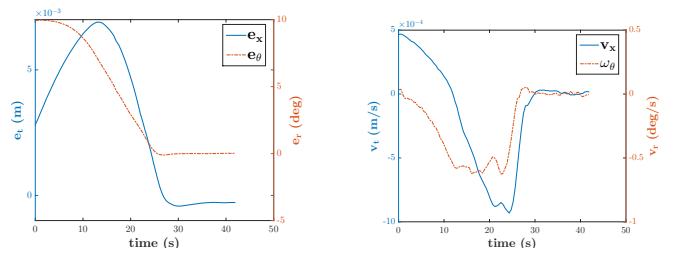


Fig. 13. (left) error decay in each DOF vs. time, and (right) velocity evolution in each DOF. (the y-translation error evolution and velocity  $v_y$  are not represented since the y-axis is controlled by the force-based control and not the visual one).

The experimental test illustrated below was reproduced

several times using photometric-based, subsampled shearlet-based and non-subsampled shearlet-based methods. To do this, 3 validation scenarios were considered: nominal conditions, under Gaussian and speckle noises. Also, for each scenario, 3 different initial positions were chosen arbitrary. Optimal control gains,  $\lambda_{Ph} = 0.03$ ,  $\lambda_{ShS} = 0.007$ , and  $\lambda_{NSHS} = 1$  are tuned for the standard photometric, our previous subsampled shearlet (ShS) and the non-subsampled shearlet (NSHS) methods, respectively. It is important to notice that each method has been computed in the more efficient way with a loop-time of 40ms, 60ms, and 200ms for the photometric, subsampled and non-subsampled shearlet methods, respectively.

Table I summarizes the obtained results in each case. It can be highlighted that the NSHS controller succeeds in each test regardless to the conditions of use. In opposition, the photometric-based method fails twice (under the speckle noise), and the ShS method fails 3 times i.e., nominal conditions with larger rotation as well as under the presence of Gaussian and speckle noises. This is certainly due to the use of a numerical interaction matrix. Concerning the accuracy, as can be expected all methods remain very accurate. Despite the fact that the computation of the non-subsampled shearlet coefficients takes more time, it remains faster than the photometric method and as fast as the subsampled shearlet one.

TABLE I

ROBUSTNESS STUDY OF THE CONTROLLER ( $e_x$  (MM),  $e_\theta$  (DEGREE),  $\Delta e_{init}$  AND  $e_f$  REPRESENT THE INITIAL AND FINAL ERRORS, RESPECTIVELY.)

$e_f$ and $e_0$	NSHS		Ph		ShS	
	$e_x$	$e_\theta$	$e_x$	$e_\theta$	$e_x$	$e_\theta$
$\Delta e_{init}$	10	8	10	8	10	8
$e_f$ Nominal	0.3	0.063	0.3	0.057	0.4	0.5
$e_f$ Gaussian	0.3	0.106	0.5	0.017	1.3	1.2
$e_f$ Speckle	0.2	0.487	0.4	0.0744	0.02	0.12
$\Delta e_{init}$	15	5	15	5	15	5
$e_f$ Nominal	0.2	0.109	0.4	0.080	0.1	0.44
$e_f$ Gaussian	0.7	0.091	0.4	0.022	0.07	0.09
$e_f$ Speckle	0.3	0.206	fail	fail	1.7	2.1
$\Delta e_{init}$	5	10	5	10	5	10
$e_f$ Nominal	0.3	0.051	0.6	0.114	fail	fail
$e_f$ Gaussian	0.9	0.143	0.3	0.051	fail	fail
$e_f$ Speckle	0.3	0.0171	fail	fail	fail	fail
average time	14.6	seconds	17	seconds	14	seconds

## VI. CONCLUSION

In this paper, was proposed a new visual servoing scheme using non-subsampled shearlet coefficients as visual signal inputs. In fact, an analytic interaction matrix form was designed by linking the time-variation of these coefficients to the robot spatial velocity. The developed controller was widely tested in different conditions of use in both simulation and experimental modes and compared to both photometric-based and subsampled shearlet-based control laws. The obtained simulation and experimental results have demonstrated the reliability of the proposed approach

in terms of convergence behavior, accuracy (**0.4mm** and **0.14°**), robustness to several US image noises (e.g., Gaussian and speckle noises), and repeatability (**9/9** successful tests against **7/9** for photometric method, and **6/9** for subsampled shearlets). Also, the proposed method demonstrated a good convergence rate (i.e., faster) during the different validation tests.

Future work will focus on the extension of the method to 6 DOF control and the consideration of other image modalities (i.e., white light camera and optical coherence tomography system). The assessment of the robustness of our approach with respect to structural evolution of the tissues will be another issue.

## ACKNOWLEDGMENT

This work is partially supported by DESWEEP project funded by the *Région de Bretagne* and from the project NEMRO (ANR-14-CE17-0013-001) funded by the ANR, France. It is also performed in the framework of the Labex ACTION (ANR-11-LABEX-01-01).

## REFERENCES

- [1] Labate, Demetrio, et al. "Sparse multidimensional representation using shearlets." *Int. Society for Optics and Photonics*, 2005.
- [2] E. King, G. Kutyniok, and W.-Q. Lim, "Image inpainting: theoretical analysis and comparison of algorithms," in *SPIE Opt. Eng. App.*. Int. Society for Optics and Photonics, pp. 885 802–885 802, 2013.
- [3] G. Easley, D. Labate, and W.-Q. Lim, "Sparse directional image representations using the discrete shearlet transform," *Applied and Computational Harmonic Analysis*, vol. 25, no. 1, pp. 25–46, 2008.
- [4] S. Yi, D. Labate, G. Easley, and H. Krim, "Edge detection and processing using shearlets," in *IEEE Int. Conf. on Image Processing*, pp. 1148–1151, 2008.
- [5] D. Gupta, R. Anand, and B. Tyagi, "Enhancement of medical ultrasound images using multiscale discrete shearlet transform based thresholding," in *Int. Symp. on Electronic Syst. Des.*, pp. 286–290, 2013.
- [6] L. A. Dufloy, A. Krupa, B. Tamadazte and N. Andreff, "Towards Ultrasound-based visual servoing using shearlet coefficients," 2016 IEEE International Conference on Robotics and Automation (ICRA), Stockholm, 2016, pp. 3420-3425.
- [7] R. Mebarki, A. Krupa, and F. Chaumette, "2D ultrasound probe complete guidance by visual servoing using image moments," *IEEE Trans. on Rob.*, vol. 26, no. 2, pp. 296–306, 2010.
- [8] M. Ourak, B. Tamadazte, O. Lehmann and N. Andreff, "Wavelets-based 6 DOF visual servoing," 2016 IEEE International Conference on Robotics and Automation (ICRA), Stockholm, 2016, pp. 3414-3419.
- [9] C. Nadeau and A. Krupa, "Intensity-based ultrasound visual servoing: modeling and validation with 2d and 3d probes," *IEEE. Trans. on Rob.*, vol. 29, no. 4, pp. 1003–1015, 2013.
- [10] A. Krupa, D. Folio, C. Novales, P. Vieyres, and T. Li, "Robotized tele-echography: an assisting visibility tool to support expert diagnostic," *IEEE Syst. J.*, pp. 1–10, 2015.
- [11] P. Abolmaesumi, S. Salcudean, W.-H. Zhu, M. Sirouspour, and S. DiMaio, "Image-guided control of a robot for medical ultrasound," *IEEE Trans. on Rob. and Aut.*, vol. 18, no. 1, pp. 11–23, 2002.
- [12] C. Collewet and E. Marchand, "Photometric Visual Servoing," *IEEE Trans. on Rob.* vol. 27, no. 4, pp. 828-834, 2011
- [13] B. Tamadazte, N. Piat, and E. Marchand, "A direct visual servoing scheme for automatic nanopositioning," *IEEE/ASME Trans. on Mech.*, vol. 17, no. 4, pp. 728–736, 2012.
- [14] F. Chaumette and S. Hutchinson, "Visual servo control. part I: basic approaches," *IEEE Rob. & Aut. Mag.*, vol. 13, no. 4, pp. 82–90 2006.
- [15] G. Kutyniok and D. Labate, "Introduction to shearlets" in *Shearlets*. Springer, pp. 1–38, 2012.
- [16] G. Kutyniok, W.Q. Lim, R. Reisenhofer. "Shearlab 3D: Faithful digital shearlet transforms based on compactly supported shearlets.", 2014.